1 Modellazione matematica del problema

Sia n il numero di partite. y_i è l'outcome di quella partita. x_i è la descrizione dalla partita i.

Sia p il numero di feature.

Vogliamo indovinare una funzione $f : \mathbb{R}^p \to \{y_0, y_1\}$

y_0 è il valore nel caso la partita è stata persa, y_1 altrimenti.

Nel nostro dataset abbiamo che $y_i = f(x_i) + \varepsilon$. dove $\varepsilon \sim \mathcal{N}(0, \sigma^2)$.

$$\mathcal{D} = \{(x_i, y_i) \mid i \in [1, n] \wedge x_i \in \mathbb{R}^p \wedge y_i \in \{y_0, y_1\}\} \tag{1}$$

Com'è composto x_i?

$$x_i = \text{StatoPoks}_i \circ \text{StatoPoksAdv}_i \circ \text{tipiPoks}_i \tag{2}$$

StatoPoks_i e StatoPoksAdv_i hanno diverse feature.

Queste feature sono la media dei vari pokemon.

Sia StatoPok_i^j lo stato di un singolo pokemon della partita i del pokemon $j \in \{1, \dots, 6\}$.

$$\text{StatoPoks}_i^f = \frac{1}{6} \sum_{j \in \{1,\dots,6\}} StatoPok_i^{j,f} \mid f \in \mathcal{F} \tag{3}$$

Dove $\mathcal{F}$ è l'insieme delle feature di un pokemon.

Nota bene: se la vita di un pokemon è 0, allora tutte le altre feature devono essere 0.

Speculare è per l'avversario.

$$\text{StatoPoksAdv}_i^f = \frac{1}{6} \sum_{j \in \{1,\dots,6\}} StatoPokAdv_i^{j,f} \mid f \in \mathcal{F} \tag{4}$$

L'unica differenza è che potremmo non poter osservare tutti i pokemon dell'avversario. Se non conosciamo il j-esimo pokemon allora possiamo inserire il pokemon medio. Ovvero la media di tutte le feature di tutti i pokemon conosciuti che possono capitare.

tipiPoks_i è il hot one encode dei tipi dei pokemon di una partita.

Per ogni pokemon possiamo fare il hot one encode dei tipi di quel pokemon. Poi vengono sommati a tutta la squadra e sottratti al hot one encode dell'avversario. Se un pokemon è morto allora non viene considerato.